

‘Kimballianen’ zien datawarehousing steeds meer gemeengoed worden

“Praktijkkennis heeft grote toegevoegde waarde”

Freek Kamst

Het ontwikkelen van een datawarehouse-omgeving is een opgave waar veel organisaties tegenop zien. Aan de ene kant lonken de voordelen van meer inzicht in de business, kostenbesparingen en dus meer winst, terwijl aan de andere kant risico's en benodigde kennis voor twijfels zorgen. Verhalen over mislukte projecten dragen ook niet bij tot een positief beeld. Dit dilemma was voor datawarehousegoeroe Ralph Kimball twee jaar geleden aanzet tot de oprichting van de Ralph Kimball University. Exclusief voor DB/M sprak Freek Kamst met twee vertegenwoordigers van de RKU, die eind mei in Nederland een seminar zullen verzorgen.

Samen met een enthousiaste groep mensen met wie hij in het verleden heeft samengewerkt, kwam Ralph Kimball op het idee van een onafhankelijke organisatie, die trainingen en seminars zou gaan verzorgen rond datawarehousing en dimensioneel modelleren. Doelstellingen van Kimballs leerinstituut zijn

Zorgvuldigheid met de businessseisen is en blijft basis voor elk datawarehouseproject

gebaseerd op kwaliteit, productonafhankelijkheid en de ervaringen vanuit gerealiseerde projecten. Vooral het laatste punt wordt gezien als een fundering onder de organisatie. “Veel sprekers en deskundigen op datawarehouse-

seminars hebben zelf nog nooit een datawarehouse ontwikkeld”, stellen Margy Ross en Warren Thornthwaite. Zij zijn beiden als deelgenoot verbonden aan de Kimballs ‘universiteit’ en hebben samen met hem diverse boeken over datawarehousing gepubliceerd. Gemiddeld nemen zij ieder jaarlijks zo'n twaalf seminars *Datawarehousing in Depth* voor hun rekening.

INLEVING

Warren Thornthwaite ontwikkelt al datawarehouses sinds 1980. “In die tijd werden dergelijke systemen nog *decision support systems* genoemd”, memoreert Thornthwaite. Margy Ross en hij werkten in de jaren tachtig beiden voor het bedrijf Metaphor Systems, toen zij daar Ralph Kimball leerden kennen. Sindsdien is hun relatie blijven bestaan en zijn zij verbonden als partner aan de organisaties van Ralph Kimball.

Op de vraag of je een datawarehouse-specialist tevens in een specifieke bedrijfstak moet zijn gespecialiseerd, zegt Thornthwaite: “Het maakt niet uit in welke ‘industrie’ je werkt, zolang je maar genoeg mensen in je team hebt die inzicht hebben in de specifieke bedrijfsvoering.” Ross vult aan: “De beginselen van een datawarehouseproject zijn voor iedere sector hetzelfde. Maar er wordt natuurlijk wel een bepaalde inleving van je verwacht in de problematiek en de terminologieën van de klant.” Zowel Thornthwaite als Ross kan een indrukwekkende lijst van projecten in diverse branches overleggen, zoals het bank- en verzekeringswezen, logistiek, telecommunicatie, de energiesector en de overheid. Maar de beide

Kimball-‘adepten’ werkten ook mee aan diens boek *The Data Warehouse Lifecycle Toolkit*, dat inmiddels min of meer fungeert als Handboek Soldaat voor data-warehouseprojecten.

60% HERSCHREVEN

Thornthwaite is momenteel bezig zelf een boek te schrijven, waarvan hij de titel nog niet wil prijsgeven. De inhoud gaat over datawarehouses in een specifieke omgeving en het voorwoord zal door Kimball worden geschreven. Ross heeft dezer dagen samen met Ralph Kimball de tweede editie van de ‘Warehouse Toolkit’ uitgebracht. Ten opzichte van het eerste boek is zestig procent herschreven, omdat door de ervaringen van de afgelopen jaren een aantal zaken beter en uitgebreider belicht kon worden. Ook geavanceerde technieken



MARGY ROSS: “DE DISCUSSIE TUSSEN RALPH KIMBALL EN BILL INMON WORDT VOORNAMELIJK IN DE MEDIA UITGEVOCHTEN. IK DENK DAT ZIJ BEIDEN GELIJK HEBBEN.”



WARREN THORNTWHAITE: "RALPHS METHODEN ZIJN UITGEBREID IN DE PRAKTIJK BEPROEFD. WAAROM Zouden WIJ ZE DAN NIET GEBRUIKEN?"

rondom het dimensioneel modelleren zijn uitgebreid in deze editie beschreven.

Op de vraag of zij in de praktijk altijd de methoden en werkwijzen van Ralph Kimball hanteren, antwoordt Thornthwaite: "Ralphs methoden komen voort uit de ervaringen van ruim zestig consultants die de afgelopen jaren met hem hebben samengewerkt en zijn dus uitgebreid in de praktijk beproefd. Waarom zouden wij ze dan niet gebruiken?"

BETAALBAAR

Temidden van de recente veranderingen in datawarehouseland valt op dat de innovatie in de hard- en software ertoe heeft bijgedragen dat datawarehouses ook voor kleinere organisaties betaalbaar zijn geworden. "Midden jaren tachtig kostte een databaseserver nog vijfhonderdduizend dollar, terwijl je nu dezelfde machinecapaciteit koopt voor ongeveer duizend dollar", aldus Thornthwaite. Een van

de laatste trends op softwaregebied is de integratie van OLAP-functionaliteit in database-systemen. Hiermee komt business intelligence al in de database beschikbaar. Ross: "Daarentegen zijn andere zaken niet veranderd. Het inventariseren en bepalen van de businessseisen en -wensen is iets wat nog steeds op dezelfde manier moet gebeuren. De zorgvuldigheid hiervan is de basis voor elk datawarehouseproject."

Is er eigenlijk nog toekomst voor datawarehousing?

Margy Ross: "Ik zie momenteel een trend naar meer realtime datawarehouses. In de eerste plaats omdat de techniek dit mogelijk maakt, maar ook omdat het management van organisaties dynamischer en sneller de juiste beslissingen moet nemen".

Thornthwaite: "Datawarehousing is nog steeds de basis voor veel managementinformatievraagstukken. Maar door de opkomst van technologieën voor CRM, e-commerce en ERP groeit de behoefte aan historisch geïntegreerde gegevens. Veel leveranciers van deze systemen zullen deze behoefte ontkennen, omdat dit een snelle implementatie in de weg zal staan. Beschikbaarheid van geïntegreerde klantinformatie is echter essentieel voor het succes van een CRM- en/of e-commercesysteem."

MINDER DESKUNDIGEN

Thornthwaite ziet de komende tijd de software om een datawarehouse te ontwikkelen steeds eenvoudiger worden. Deze software zal er voornamelijk toe bijdragen dat een datawarehouse omgeving sneller te realiseren is met minder deskundigen. "Ik maak nu al mee dat in mijn klas één of twee medewerkers van een

organisatie verantwoordelijk zijn voor de bouw van een datawarehouse. In de nabije toekomst zie ik dan ook binnen kleinere organisaties dat de databasebeheerder ook diegene is die het warehouse ontwikkelt en beheert". Hierdoor zal geïntegreerde informatievoorziening steeds meer beschikbaar komen voor kleinere organisaties.

Zal de 'eeuwige strijd' tussen Bill Inmon en Ralph Kimball blijven voortduren?

Margy Ross: "De discussie tussen Ralph en Bill wordt voornamelijk in de media uitgevochten. Ik denk dat zij beiden gelijk hebben. Maar de uitleg van Ralph over datamarts wordt niet altijd begrepen. Bill ziet datamarts als onafhankelijke subjecten, terwijl Ralph de datamart als positioneert als een busstructuur, waarin je logische subjecten kunt inpluggen. Net zoals je in een pc componenten op de bus aansluit, kun

"Het management moet steeds sneller de juiste beslissingen nemen"

je in een logische datawarehouseomgeving dit ook doen met datamarts. De dimensies zijn dan de bus en de feiten zijn als het ware de componenten die je inpluigt. Zo is Ralphs architectuur gebaseerd op een modulaire structuur, waarmee je iteratief logische subjecten kan toevoegen. Deze subjecten noemt Ralph datamarts, die na een aantal iteraties vanzelf uitgroeien tot een datawarehouse."

DATAMART 'UIT DE DOOS'

Het *datamart in a box*-concept is een fenomeen dat sinds een aantal jaren door diverse leveranciers tot de nieuwe generatie datawarehousetoepassingen wordt gerekend. Het concept gaat uit van standaard tools en een op sectoren toegesneden standaard datamodel. Thornthwaite en Ross melden dat dit concept steeds succesvoller in de Verenigde Staten wordt toegepast. Alleen worden standaard datamodellen steeds minder gebruikt. Men stapt nu meer over op datamarts die standaard voor specifieke toepassingen zijn ontworpen. Zo komen op ERP-gebied allerlei op maat ont-

Quest for Knowledge

In de Benelux vertegenwoordigt Quest for Knowledge (Q4K) de Ralph Kimball University (RKU). Q4K heeft al geruime tijd een samenwerkingsverband met de naamgever van dit instituut, omdat ook zijn persoonlijke seminars door deze organisatie worden georganiseerd. Zelf verzorgt Q4K ook onderwijs volgens het RKU-concept, waarbij productionafhankelijke datawarehousetrainingen worden gehouden door ervaren specialisten. Eind mei 2002 komen Warren Thornthwaite en Margy Ross naar Nederland voor een seminar om de hechtheid van de RKU/Q4K-relatie nog eens te onderstrepen en voor de promotie van het nieuwe boek van Ralph Kimball en Margy Ross, *The Data Warehouse Toolkit Second Edition*. In dit interview met Ross en Thornthwaite wordt ingegaan op de actualiteit en de toekomst van datawarehouses.

wikkelde datamarts op de markt. Dit komt doordat bij deze toepassingen tot nog toe weinig aandacht werd besteed aan het aspect van informatievoorziening. Maar ook bij andere standaard toepassingen is het rapportage- en informatievoorzieningsgedeelte meestal 'vergeten'.

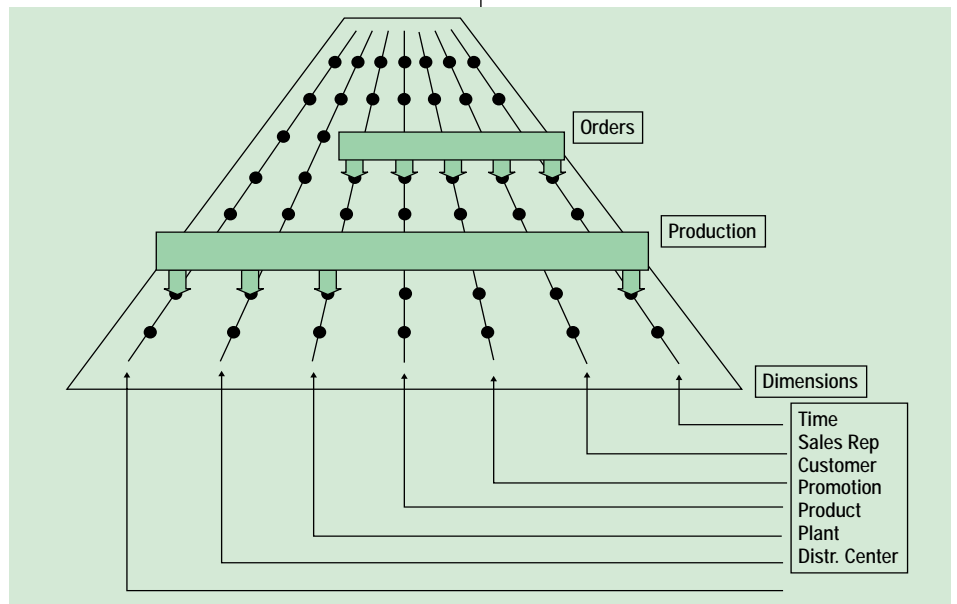
De behoefte aan informatie ontstaat pas na de implementatie, en meestal levert de leverancier zelf of een *third party* daarvoor vervolgens een datamartoplossing. Margy Ross: "De standaard datamodelen die diverse leveranciers al geruime tijd aanbieden, blijken in de praktijk sterk af te wijken van de eigenlijke informatiebehoefte. Aanpassing kost evenveel tijd als het opnieuw ontwikkelen hiervan. Bovendien wil de klant zich onderscheiden in de markt en een voorsprong op de concurrentie behalen. Als iedereen dezelfde standaard oplossing gebruikt, is dit concurrentievoordeel niet meer relevant."

VALKUILEN

Datawarehouseprojecten zijn altijd als uiterst risicovol bestempeld. Welke drie valkuilen in zo'n project vragen volgens de beide Kimball-deskundigen de meeste aandacht?

- *Bouwen van een datawarehouse zonder de eisen en wensen te beschouwen vanuit de business.* Als een project wordt geïnitieerd vanuit de ICT-afdeling -bijvoorbeeld om een geïntegreerd gegevensbeeld aan te bieden- is het project op voorhand gedoemd te mislukken. Zorg er altijd voor dat de business nauw betrokken is bij opzet en ontwikkeling.
- *Ontbreken van een sponsor.* Maar al te vaak gaat een datawarehouseproject te laag in de organisatie van start. Heeft het initiatief niet voldoende budget of wordt het budget via andere kanalen losgemaakt, dan is er meestal geen goede financiële fundering voor een zo'n project. Zorg voor een realistische budgettering en zorg dat het wordt gedragen door het management.
- *Een datawarehouse gebouwd zonder gebruikersfocus.* Wie de gebruikers pas bij de implementatiefase confronteert met het warehouse is te laat. Zorg ervoor dat zij voldoende zijn vertegenwoordigd en gebruik ze als klankbord.

Waarom bestaat veertig procent van uw klas-



DE DATAMART-BUSSTRUCTUUR VOLGENS RALPH KIMBALL.

sen uit ervaren datawarehousespecialisten?

Margy Ross: "Onze klassen zijn gebaseerd op meer dan vijftien jaar datawarehouse-ervaring. Voor diegenen die voor het eerst in aanraking komen met datawarehousing hebben de verhalen over valkuilen en de praktijkervaringen grote toegevoegde waarde bij de start van een project. Voor specialisten die al een datawarehouseproject achter de rug hebben, is het een toetsing van de gedane werkzaamheden. En voor er nog middenin zit, een bijsturing in een lopend project."

TOEGANKELIJKER

Samenvattend kan men stellen dat de ontwikkeling van een datawarehouse-omgeving nog steeds een kennisintensieve bezigheid is, maar dat op termijn dergelijke projecten toegankelijker zullen worden voor beheerders en

business-gebruikers. Kijkend naar de productlevenscyclus van een datawarehouse op dit moment, zien we dat de top van de curve net is bereikt. Met de opkomst van CRM-, e-commerce- en ERP-toepassingen zal de informatiecomponent steeds belangrijker worden. De samensmelting van alle gegevens in een geïntegreerde omgeving zal uiteindelijk toch buiten de toepassingen zelf blijven plaatsvinden, tenzij men één leverancier vindt die de gehele informatievoorziening binnen een organisatie kan invullen. De vraag blijft of een klant het risico van een dergelijke afhankelijkheid wil nemen. Voorlopig blijven er met betrekking tot datawarehousing nog voldoende uitdagingen over. Een daling in de curve is nog niet in zicht.

Freek Kamst (freek.kamst@newcom.nl) is consultant bij Newcom Information Systems.

Competitive intelligence

Een nieuwe term die de laatste tijd steeds meer gehoord wordt is *competitive intelligence*. Warren Thornthwaite hierover: "Competitive intelligence speelt een rol in twee gebieden. In de eerste plaats is er de opkomst van tools en methodieken om niet-gestructureerde gegevens te categoriseren en beschikbaar te maken. Dit is het werkterrein van knowledge management en metadatabeheer. Door interne en externe gegevens gestructureerd beschikbaar te maken, ontstaan inzichten die kostenbesparend en concurrerend zijn. Ten tweede zie je veel organisaties gegevens verzamelen over bepaalde branches, om die vervolgens op de markt aan te bieden. Dit fenomeen is momenteel in de Verenigde Staten enorm in opkomst. Zowel de overheid als het bedrijfsleven ziet veel toekomst in het leveren van deze gegevens."